Supercoiled DNA

In many forms of DNA the double helix itself forms a helix. This supercoiling, which has important biological consequences, is best described and analyzed by means of a simple mathematical model

by William R. Bauer, F. H. C. Crick and James H. White

DNA, the primary genetic material of most organisms, is usually visualized as a double helix in which two chains of complementary nucleotides (the subunits whose sequence constitutes the genetic message) wind around a straight common axis. It is now clear, however, that often the axis of the double helix is not linear but curved. Indeed, the double helix can wind in space to form a new helix of a higher order, in which case it is said to be supercoiled. It appears that a large proportion of the known DNA's exhibit some form of supercoiling in at least one stage of their life cycle. An appreciation of this structural feature and its consequences is therefore essential to a complete understanding of the biology of DNA.

Affecting DNA's in a wide range of sizes and shapes (including some that are not organized into a double helix), supercoiling takes a variety of forms. For example, in the chromatin (DNA complexed with protein) of higher organisms the DNA is wound around a core of protein to form a left-handed solenoidal superhelix. Here we shall be concerned mainly with a different type of supercoiling in which no protein core is needed. It is the supercoiling of closed circular DNA: double-helix molecules in which each polynucleotide chain forms an unbroken loop.

The existence of these twisted rings of DNA was first postulated to explain a surprising discovery about the DNA of the polyoma virus: a small virus that causes tumors in the mouse. In 1963 it was reported that when this DNA is suspended in a solvent and spun in a centrifuge, it resolves into three components distinguishable by the velocity at which they move through the solvent. Further investigation revealed that these components, labeled I, II and III in order of their decreasing velocity of sedimentation, do not differ in molecular weight. Hence there had to be some variation in the molecular compactness of the three species accounting for the difference in their velocities. In other words, the molecules had to have different shapes.

Viewing the components by means of electron microscopy indicated that this was indeed the case. Although the molecules of component III were clearly linear, those of component I and component II appeared to lack ends. The revelation that components I and II were circular, however, did not solve the problem. There was still no clue to what distinguished the two components. Then in 1965 Jerome Vinograd of the California Institute of Technology put forward an ingenious proposal that solved the mystery and introduced the concept of supercoiling in DNA.

Vinograd suggested that a closed circular molecule of DNA (a circular molecule in which both polynucleotide chains are completely intact) is usually underwound compared with a linear piece of DNA of the same length. In other words, in the closed circular molecules of component I there are fewer helical turns than there are in the molecules of either component II or component III. Component I behaves as though before the ends of the double helix were joined a number of 360-degree twists had been made in the direction opposite to that of the helix's normal winding.

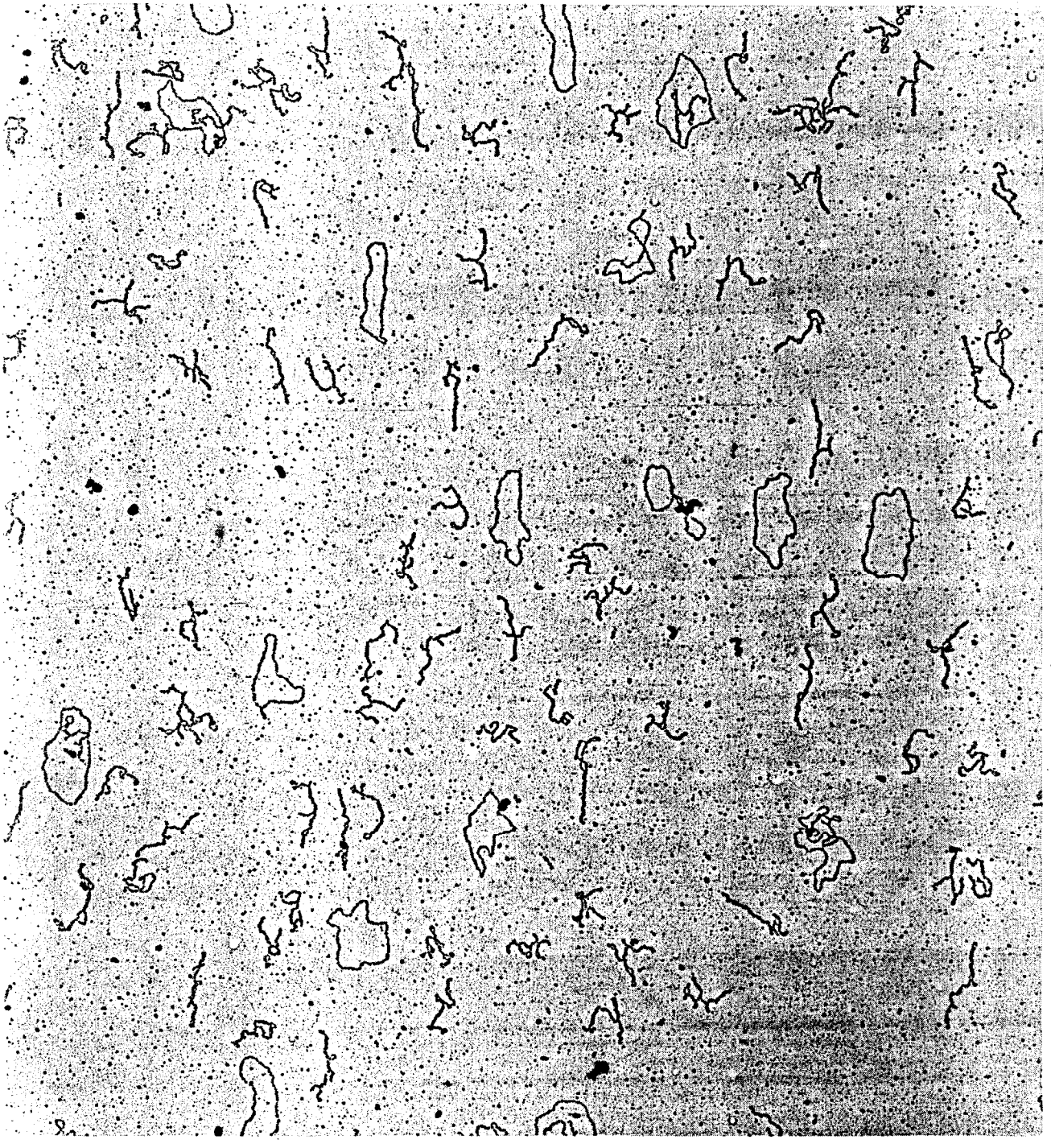
The difference between components I and II of polyoma DNA, then, is that the circular molecules of component II are "nicked": in each molecule there is at least one nick, or break, in one of the polynucleotide chains. In contrast, circular molecules of component I are closed and underwound. Because the forces that stabilize the double helix are strong the closed circular molecules resist such underwinding, and for reasons we shall discuss they compensate by forming supercoils. Hence the closed circular molecules of component I of polyoma DNA are more compact than the nicked circular molecules of component II and therefore have a higher sedimentation velocity.

Supercoiling is a widespread phenomenon, characterizing many of the medically and biologically most interesting DNA's. An entire class of DNA tumor viruses, including the polyomas and the human papilloma (wart) viruses, contain such DNA, and the DNA of mitochondria (the energy-transducing organelles) of human and other animal cells is supercoiled. Animal cells have also been found to contain a vast array of very small supercoiled DNA molecules whose function has not been determined. It is particularly striking that the majority of the known smaller genomes (sets of genes) fall in this category, including genetic factors for fertility and drug resistance. Moreover, the phenomenon of integration, in which a small piece of DNA is physically inserted into a larger DNA molecule, appears to require that the integrated element be supercoiled. Thus the vectors, or vehicles, for DNA in genetic engineering are of this type. Perhaps most remarkable is the fact that RNA tumor viruses, including the cancer-causing Epstein-Barr virus, often employ the enzyme reverse transcriptase to make DNA copies that become supercoiled in the course of infection. It should also be noted that supercoiling has been found in molecules with as few as 350 nucleotides in each polynucleotide chain and with as many as 1,750,000.

Virtually every physical, chemical and biological property of DNA—its transcription, hydrodynamic behavior, energetics, enzymology and so on—are affected by closed circularity and the deformations associated with supercoiling. Understanding the mechanism of supercoiling and the consequences of this structural feature for DNA, however, presents problems of considerable mathematical complexity. Fortunately there are two branches of mathematics that offer substantial help in this effort: topology, which studies the properties of structures that remain unchanged when the structures are deformed, and differential geometry, which applies the methods of the differential calculus to the study of curves and surfaces. In what follows we shall first describe a mathematical model of closed circular DNA and then discuss the implications of the

model for real DNA. We shall also take up some of the methods that have been developed for measuring supercoiling and its effects.

Consider first the physical structure of the double helix: it is formed by two sugar-phosphate backbones to which the nucleotide bases of the DNA are attached. The bases on the opposing backbones are paired, with about 10 base pairs for each turn of the helix. (The exact number depends on the particular configuration of the DNA molecule.) To study supercoiling mathematically it is most convenient to construct a model in which the structure is represented as a narrow twisted ribbon of infinitesimal thickness. The most obvious way to design such a model is to specify that the edges of the ribbon follow the sugar-phosphate backbones of the DNA. Carrying out the construction in the simple case of a linear double-helix molecule,



MANY SUPERCOILED DNA MOLECULES appear in this electron micrograph. Each molecule is a loop of the DNA double helix. Where the loops are open one of the strands of the double helix is "nicked," or broken, so that the loop is relaxed, or not supercoiled. Where the loops are not open but kinked they are un-nicked. In such molecules the DNA usually has fewer helical turns than a linear molecule of the same length. In un-nicked loops the stress introduced by this underwinding is compensated for by supercoiling. These molecules are plasmids extracted from the bacterium *Escherichia coli* and separated from other *E. coli* DNA molecules by centrifugation. They are enlarged some 25,000 diameters. The micrograph was made by Gary Cohen of the State University of New York at Stony Brook.

however, demonstrates that this choice is unsatisfactory: the axis of the model (the line halfway between the edges of the ribbon) winds around the straight axis of the double helix. For our purposes it is preferable to choose a model whose axis coincides with that of the double helix. In addition we specify that the ribbon must always lie perpendicular to the pseudodyads, or twofold axes of rotation, that are distributed along the double helix [*see illustration below*]. (There is one dyad axis associated with each nucleotide pair and one associated with the space between successive pairs; these axes are all perpendicular to the axis of the double helix, and their location is independent of the sequence of bases in the molecule.)

This ribbon model follows the axis of the DNA double helix and twists as the two chains of the molecule twist around that axis. In addition, because the sequences of atoms in the two polynucleotide chains run in opposite directions the edges of the ribbon will be assigned opposite orientations. (It does not matter what directions are chosen for the two edges as long as they are opposite.) This model can be analyzed mathematically in a number of different ways, but we shall be most concerned here with the relation between the oppositely directed edges of the ribbon.

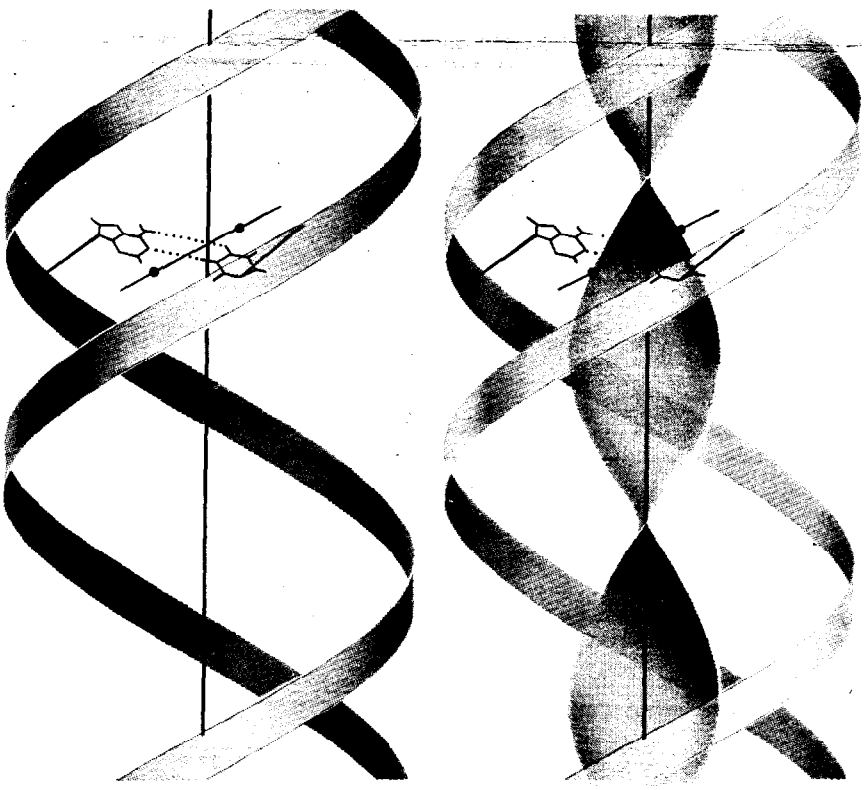
When the ends of a ribbon are joined, each edge describes a closed curve in three-dimensional space. Furthermore, when the ribbon represents a closed circular molecule of DNA, a number of 360-degree twists are introduced before the ends of the ribbon are joined, and so the two curves described by its edges are linked. In other words, it is impossible to separate the curves without "cutting" one of them. If each loop in a linked pair represents a covalently bonded molecule, as is the case with the two polynucleotide chains of the double helix, the two are said to be joined by a topological bond. This is a peculiar type of bond in that although no part of one molecule is covalently joined to any part of the other, it is nonetheless necessary to break a covalent bond in order to separate the two. (The concept of topological bonding was first introduced into chemistry by Edel Wasserman and his colleagues at Bell Laboratories.)

In mathematical terms the linking of two closed curves is a topological property: no matter how the curves are deformed (pulled, twisted and so on), as long as neither one is broken they will remain linked in exactly the same way. It will be useful, then, to assign a numerical value that describes the way the loop formed by one edge of the ribbon representing a molecule of DNA is linked to the loop formed by the other edge.

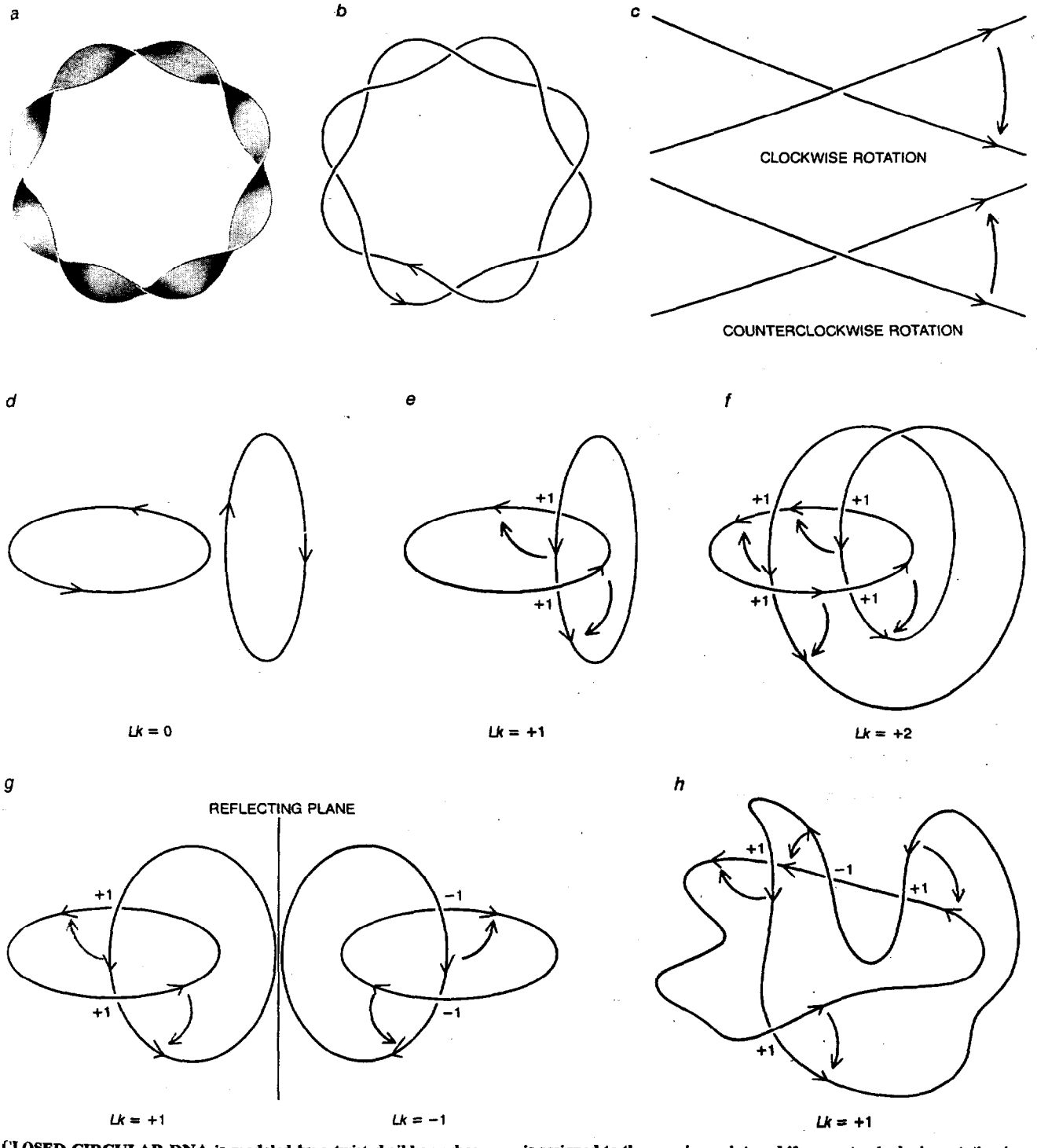
It seems reasonable to define such a linking number Lk so that its value will be 0 for a pair of unlinked closed curves, 1 for a curve that loops through another just once, 2 for a curve that loops through another twice and so on. Since the two edges of the ribbon have been given orientations, however, it turns out that a more useful definition of a linking number can be made in which either a plus or a minus sign is assigned to the number depending on the orientation of the curves. There are a number of ways to compute such a linking number for a particular configuration of curves, and one of the most convenient is to examine all the points in a projection, or two-dimensional representation, of the configuration where a piece of one of the curves crosses over a piece of the other [*see illustration on opposite page*]. Each of these crossing points can be assigned an index number of $+1$ or -1, according to the direction in which the top piece must be rotated so that it coincides with the bottom piece. Adding up the index numbers of all the crossing points and dividing by 2 (the number of linked curves) gives the linking number Lk.

Thus the linking number Lk is defined as a signed integer that describes a property of two closed curves in space. To separate a pair of curves without actually cutting them the value of Lk must be 0 (although the converse is not always true). If the curves in question are the edges of a closed ribbon with N turns in it, their linking number is $+N$ or $-N$, depending on the direction of turning. Moreover, as long as the ends of the ribbon remain joined the number will remain unchanged when the ribbon is deformed. Notice that although the edges of the ribbon model of DNA were not chosen to coincide with the sugar-phosphate backbones of the double helix the linking number of the ribbon will be exactly that of the backbones.

Therefore for a molecule of relaxed closed circular DNA 5,000 base pairs long, with 10 base pairs for each turn of the helix, the linking number will be $+500$. By convention the number is positive because the DNA double helix is right-handed. (A closed circular DNA molecule is said to be relaxed if its axis lies entirely in a plane. As we shall relate, there are many ways in which



MODEL OF SUPERCOILED DNA makes use of two structural features of the double helix (*left*): the axis (*thick colored line*) around which the two sugar-phosphate backbones of the molecule wind and the series of "pseudodyad" axes defined by the nucleotide bases attached to the backbones. The bases, whose particular sequence constitutes the genetic message, are paired, and there is one dyad axis associated with each base pair and one with the space between successive base pairs. (Here only one base pair is shown.) The dyad axis (*thin colored line*) defined by each base pair is the line perpendicular to the helical axis about which either one of the bases can be rotated into the position of its oppositely oriented mate. The model constructed with respect to these features (*right*) is a narrow, infinitesimally thin ribbon whose axis, or centerline, follows the axis of the double helix and whose surface always lies perpendicular to the dyad axes defined by the base pairs of the double helix. This ribbon bends as the axis of the DNA molecule bends, and it also twists as its two polynucleotide chains twist around each other.



CLOSED CIRCULAR DNA is modeled by a twisted ribbon whose ends have been joined (*a*). Since the sequences of atoms in the two chains of the double helix run in opposite directions, it is also convenient to assign the edges of the ribbon opposite directions. When the edges are viewed as directed closed curves in three-dimensional space (*b*), they are found to be mathematically linked; in other words, there is no way to separate them without breaking one or the other. This relation can be described mathematically by a linking number *Lk* whose magnitude expresses the number of times one curve is linked through the loop of the other and whose sign depends on the way the curves are labeled. One way to compute the value of this quantity is to examine a projection, or two-dimensional representation, of the two curves and to each point where one curve crosses over the other (but not where a curve crosses over itself) assign an index number according to the following rule: If a clockwise rotation is required to move the top piece so that it coincides with the bottom piece, then +1 is assigned to the crossing point, and if a counterclockwise rotation is required, then −1 is assigned (*c*). (This convention is the reverse of the one normally employed in mathematics, but it ensures that the linking number of right-handed closed circular DNA will be positive.) *Lk* is then calculated by adding up the index numbers and dividing by 2 (the number of curves). The linking number obtained in this way is a signed integer equal to 0 if the two curves are unlinked (*d*), to +1 or −1 if one curve links through the other just once (*e*), to +2 or −2 if one curve links through the other twice (*f*), and so on. The sign of the number will change if the orientation of either one of the curves is changed or if the pair of curves is viewed in a mirror (*g*). The value of the linking number remains the same no matter how the two curves are deformed (*h*), and so since the linking number of a twisted ribbon is equal to the linking number of the sugar-phosphate backbones of the DNA molecule it models, the linking number expresses an important constraint on the possible supercoiled structure of the DNA.

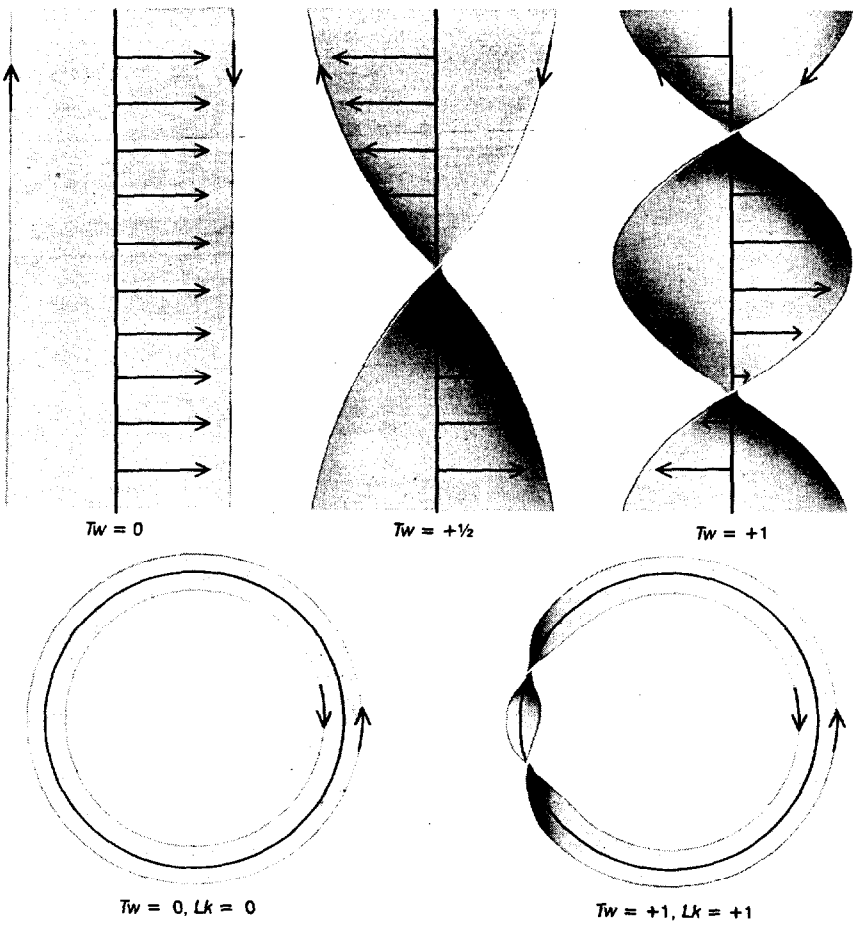
naturally supercoiled molecules can be made to relax. A nicked circular molecule is relaxed in its natural state, but since its polynucleotide chains are not both intact, strictly speaking it has no linking number, the quantity that is our principal concern here.)

Another way to analyze the ribbon model of DNA is by looking not at the relation between its edges but at the way the ribbon twists. For a ribbon whose axis follows a straight line the idea of a numerical value expressing twist is intuitively obvious. Here we shall adopt the convention that a right-handed twist of 360 degrees has a value of $+1$ and a left-handed twist has a value of -1. The definition of twist is less obvious, however, for a ribbon whose axis is not straight. Perhaps the best way to understand this concept is to imagine a small arrow placed perpendicular to the axis of the ribbon, pointing to one of its edges [*see illustration below*]. As the arrow is moved along the twisting ribbon it rotates about the axis, and the twist of the ribbon can be defined as the integral of the arrow's angular rate of rotation with respect to the arc length of the axis curve.

In the special case where the axis of the ribbon is confined to a plane this value can be measured simply as the number of rotations the arrow completes about the axis as it is moved along the ribbon. For example, when the ribbon models a closed circular piece of DNA 5,000 base pairs long that is relaxed (that is, its axis lies in a plane), the arrow makes one complete rotation for every turn of the double helix, and so the total twist Tw equals $+500$, with the plus sign arising once again because the double helix is right-handed.

For a relaxed circular piece of DNA 5,000 base pairs long, then, both the linking number and the twist are equal to $+500$. From this example one might well assume that linking number is just another way of expressing twist, but that is not the case. Indeed, it is particularly important to understand the distinction between these two quantities. To begin with, linking is a topological property, whereas twist is geometrical: if a ribbon is deformed, its twist may be altered. Moreover, to compute the linking number (which is always an integer) the ribbon must be considered as a whole. On the other hand, twist (which might not be an integer) can be considered locally, and the twist values of individual sections can be summed to obtain the total for the ribbon.

It is important to note that this ribbon model does not come close to representing the detailed mechanical properties of DNA, not least because the mathematical ribbon is assumed to be of infinitesimal thickness. A real ribbon (with finite thickness) subjected to the helical manipulations we are concerned with would experience numerous local deformations that are not taken into account here. The ribbon model, however, was designed to emphasize the linking number of DNA, which because it is quantized is not affected by such small deformations.



TWISTING OF A RIBBON can be assigned a numerical value by placing a small arrow on the ribbon perpendicular to its axis and pointing to one of its edges. As the arrow moves along the ribbon it rotates about the axis, and the magnitude of the total twist Tw can be defined as the integral of the angular rate of this rotation with respect to the arc length of the curve described by the axis. Whether this quantity is positive or negative depends on whether the rotation of the arrow about the axis is respectively right-handed or left-handed. When the axis of the ribbon lies entirely in a plane as is shown here, then the total twist is easily computed, being equal to the number of rotations the arrow makes about the axis. The twist can be computed separately for different parts of the ribbon and can then be summed to obtain the total value.

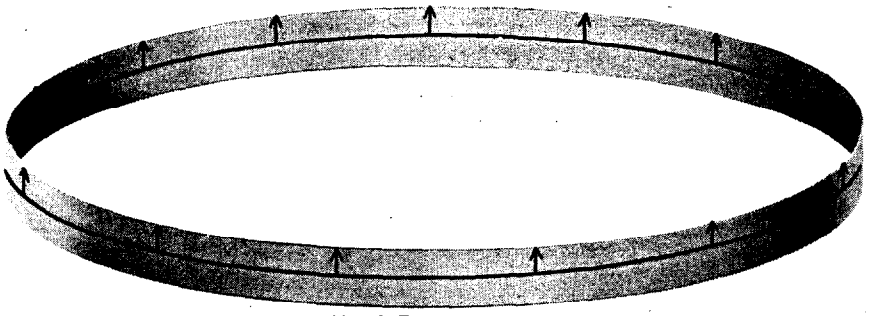
The realization that linking and twisting are distinct properties raises another question: Is there a geometrical significance to the difference between these properties, that is, to the difference between the linking number of a ribbon and its total twist? Quite fortuitously at about the same time that biochemists were first studying circular DNA mathematicians were independently taking a look at linking and twisting in ribbons. In 1968 one of us (White) proved that the linking number of a ribbon and its total twist differ by a quantity that depends exclusively on the curve of the axis of the ribbon. (This quantity is well known to mathematicians as the Gauss integral of the axis curve.) In other words, assume that the axes of two closed ribbons follow the same curve in three-dimensional space; then even if the ribbons themselves turn and twist in entirely dissimilar ways, their values of linking number and total twist will differ by exactly the same amount.

At about the same time Vinograd, not knowing of this result, asked F. Brock Fuller, a mathematician at Cal Tech, to tackle the mathematics of supercoiling. Fuller, while working on the relation between linking and twist, suggested the picturesque name writhing number for the quantity by which the two differ. Thus for a closed ribbon in three-dimensional space the writhing number Wr equals the difference between the linking number Lk and the total twist Tw, or $Wr = Lk - Tw$. The writhing number of a ribbon is a powerful quantity whose value generally changes if the axis of the ribbon is deformed in space. Hence

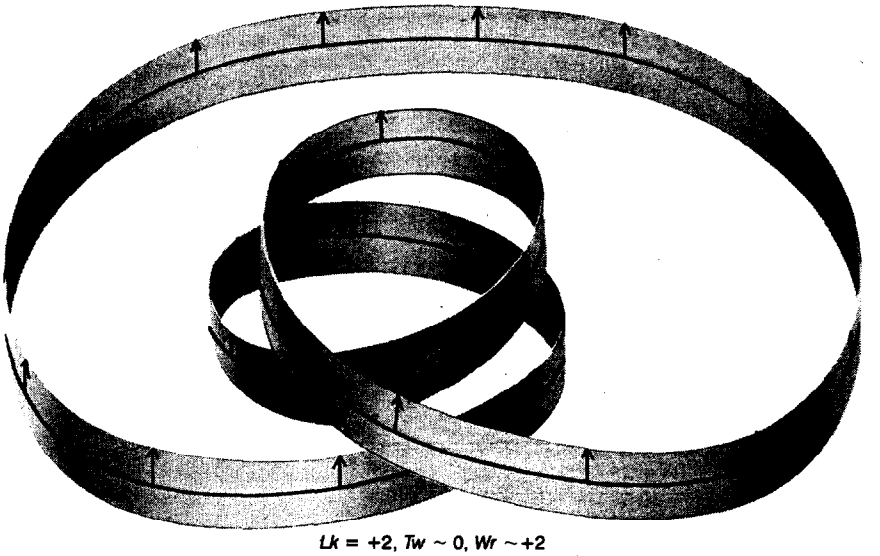
writhing, like twisting, is not a topological property of the ribbon but a geometrical one.

The writhing number can be obtained by computing the Gauss integral, but it is generally far easier to calculate it by evaluating the linking number and the total twist of the ribbon in question and then taking their difference. It is only in certain special cases that it is convenient to compute *Wr* directly. For example, if the axis of a ribbon lies entirely in a plane or entirely on the surface of a sphere, then it can be shown that *Wr* is zero. Substituting this value into the equation $Wr = Lk - Tw$ gives $Lk = Tw$, which explains why in the example of the relaxed closed circular molecule of DNA both *Tw* and *Lk* were found to be +500. Now consider what happens if the axis of this DNA molecule is made to writhe in such a way that its writhing number is no longer zero. (There is no intuitive way to estimate the writhing number of a curve. A ribbon that writhes in the ordinary sense of the word may turn out to have an overall writhing number of zero.) When the writhing number of the molecule is made to change, the linking number remains the same (it can be altered only if one of the backbones of the double helix is broken) and so the twist must change. It is this relation that underlies the phenomenon of supercoiling.

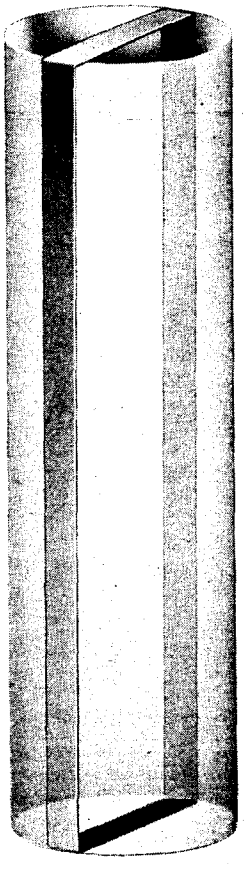
A ribbon's linking number, total twist and writhing number do not depend on the ribbon's location or orientation in space. They are also independent of scale, but if one axis of space is inverted (as is the case when the ribbon is reflected in a mirror) or three axes are inverted (as is the case when the ribbon is inverted through a point), then the sign of all three quantities is changed. On the other hand, if any two axes are invert-



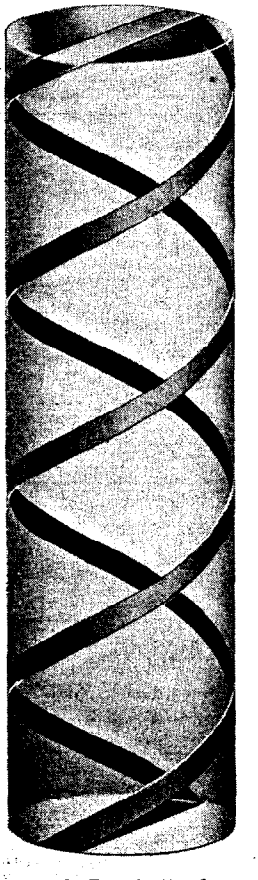
$Lk = 0, Tw = 0, Wr = 0$



$Lk = +2, Tw \sim 0, Wr \sim +2$

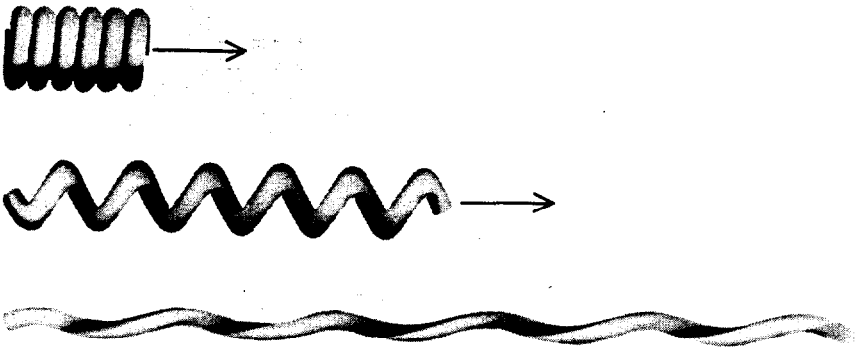


$Lk = 0, Tw = 0, Wr = 0$

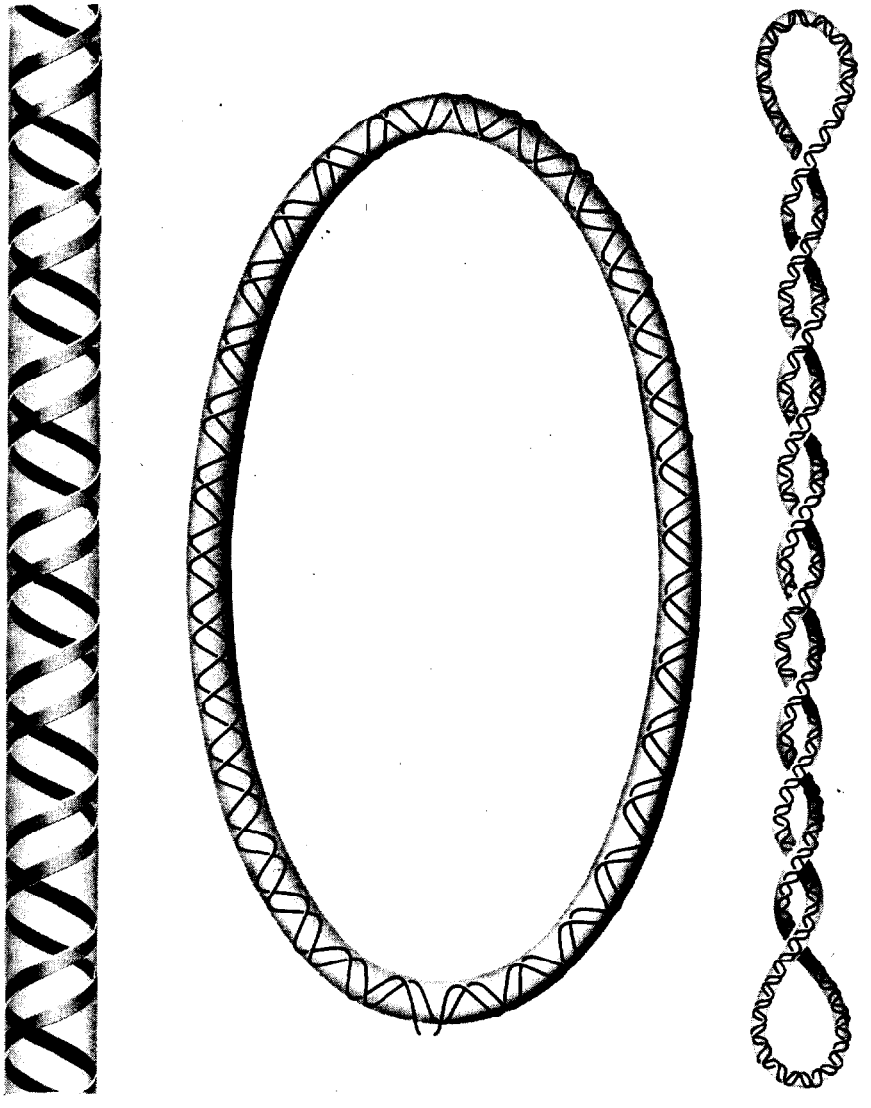


$Lk = 0, Tw > 0, Wr < 0$

LINKING NUMBER and total twist of a ribbon are equal when the axis of the ribbon lies entirely in a plane as is shown at the top. That these two quantities do not necessarily have the same value is shown by the configuration in the middle, in which the linking number *Lk* is a positive integer and yet the total twist *Tw* is close to zero. The linking number and total twist of a ribbon differ by a quantity called the writhing number *Wr*, that is, $Wr = Lk - Tw$. The writhing number depends exclusively on the curve of the axis of the ribbon. In the ribbon wound around the cylinder shown at the bottom left *Lk*, *Tw* and *Wr* are all zero. Rotating the top of the cylinder counterclockwise about its central axis through two full turns generates the right-handed interwound helical ribbon shown at the bottom right, in which *Lk* is clearly unchanged but (neglecting end effects) *Tw* equals $+4\sin\alpha$ and therefore *Wr* equals $-4\sin\alpha$. (In the first case the total twist *Tw* was made to be zero, and so the final configuration is a simplified model of supercoiled DNA, *omitting the intrinsic twist* that arises from the turns of the double helix.)



RELATION BETWEEN WRITHING AND TWIST is demonstrated with a coiled telephone wire. When the wire is in a relaxed state (*top*), its twist is small and its writhing is substantial: the axis of the wire traces a helix in three-dimensional space. When the wire is stretched out (*middle*) so that its axis is nearly straight (*bottom*), its writhing is small and its twist is large.



BASIC CONFIGURATIONS OF THE DOUBLE HELIX include a linear form (*left*), a circular form that is nicked and therefore relaxed (*middle*) and a circular form that is closed (unnicked) and therefore supercoiled (*right*). (The three molecules shown here are drawn to different scales.) The supercoiling of the closed circular molecule into an interwound superhelix can be understood in terms of the relation between linking, writhing and twist. Because molecule is underwound it has a deficit in linking number compared with a relaxed molecule of the same size. It compensates by writhing and by twisting and bending, satisfying equation $Wr = Lk - Tw$.

ed, as happens if one looks into an ordinary microscope (that is, not a dissecting one), their signs are unchanged. There are two other cases of practical interest. Making a contact print of a projection of the ribbon changes the signs of *Lk*, *Tw* and *Wr*. On the other hand, if the negative of a photograph of this projection is placed correctly in an enlarger, the resulting print will show a ribbon with the signs of all three quantities unchanged. Conversely, if the negative is turned over in the enlarger, the signs of *Lk*, *Tw* and *Wr* in the resulting print will be reversed.

In short, any operation that turns a right-handed screw into a left-handed one without introducing other distortions will change the sign of the linking number, the total twist and the writhing number. One of us (White) and Thomas F. Banchoff of Brown University have shown that there is one other special mathematical operation that changes the sign of these quantities but leaves their magnitude unaltered: inverting the ribbon through a sphere. This result explains why the writhing number of a ribbon whose axis lies on the surface of a sphere is zero. Under this operation the closed curve described by the axis is transformed into itself.

Although the equation $Wr = Lk - Tw$ demonstrates that linking and twist are mathematically distinct, the physical difference between these quantities may not yet be evident. It may be helpful, then, to consider what happens when a mathematical ribbon is wound around a cylinder in such a way that its surface is always flat against the cylinder [*see illustration on preceding page*]. (Note that it is only because a mathematical ribbon has infinitesimal thickness that it can be made to lie flat against the surface of a cylinder in this way.) We shall call the pitch angle of the helix described by this ribbon α. In other words, α is the angle at which each turn of the helix inclines away from the horizontal, so that when α is small, the helix is shallow, and when α is large, the helix is steep.

Now assume the ribbon is wrapped around the cylinder N times before its ends are joined in the most straightforward way. Then if end effects are ignored, it can be demonstrated that the linking number of the ribbon Lk equals N, whereas the total twist Tw equals $N\sin\alpha$. Therefore when the helix is stretched out so that the pitch angle α increases, the number of turns and thus the linking number remain the same, but the twist goes from a small value to a large one, clearly demonstrating the difference between linking and twist.

Moreover, since a ribbon's writhing number is defined as the difference between its linking number and its total twist, the value of Wr for this ribbon is $N - N\sin\alpha$, or $N(1 - \sin\alpha)$. As this formula indicates, when α is small and the

twist is small, the writhing is substantial, but when α is large and the twist is large, the writhing is minimal. The relation can be easily observed in a coiled telephone wire: when such a wire is unstressed, it assumes a highly writhed form with little twist; when its ends are pulled out, a highly twisted form that writhes only slightly is obtained.

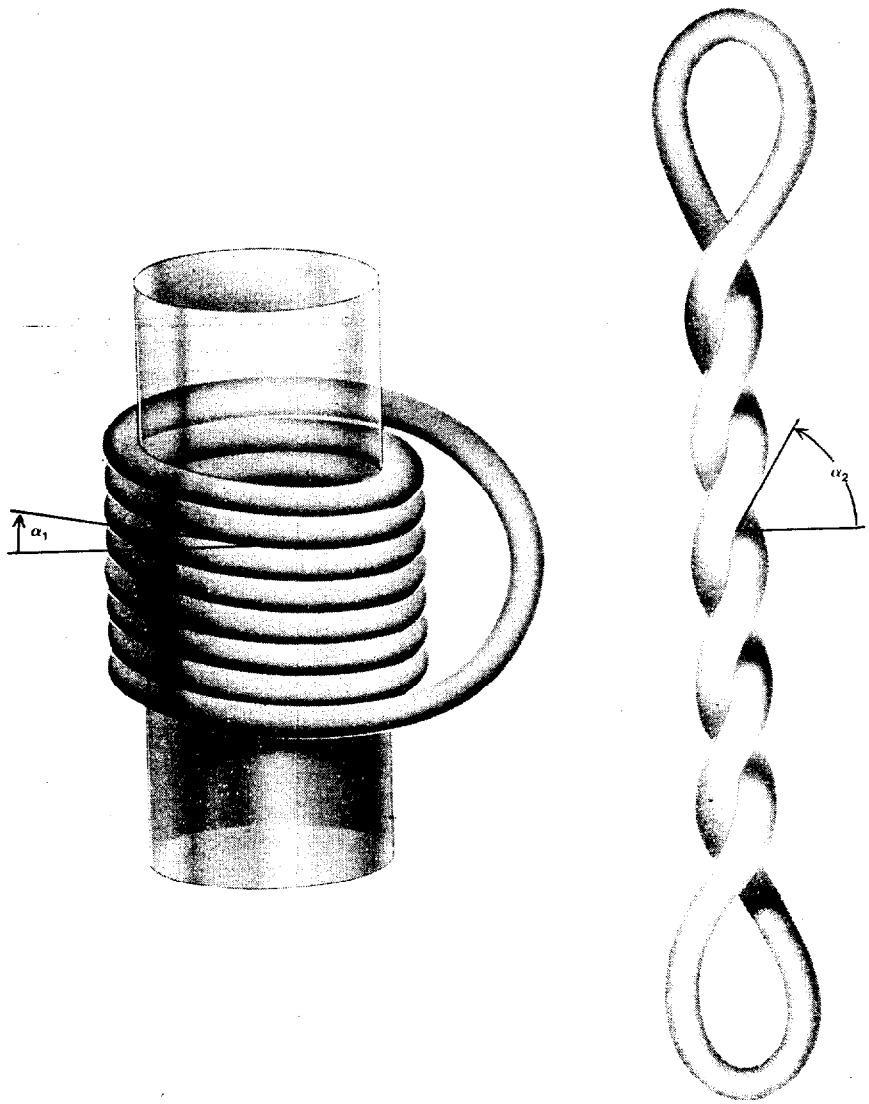
Consider how these findings apply to a real DNA, say to the polyoma-virus DNA. Remember that this DNA can be resolved through sedimentation into three components: I and II, which are circular, and III, which is linear. It can be determined experimentally that the average linking number for a population of relaxed circular molecules of this DNA is about +500. (Even in a highly purified preparation not all molecules are relaxed or supercoiled to the same extent and so the average linking number of a population is not necessarily an integer.) On the other hand, for a population of closed circular molecules (the supercoiled molecules that make up component I) the average linking number is about +475. Just as Vinograd predicted, the closed circular molecules of the polyoma-virus DNA are underwound, having a deficit in their winding number of about 25. This finding suggests a way to define supercoiling. It is equal to ΔLk, the difference between the linking number of a molecule in the natural closed circular state and the linking number of the same molecule in the relaxed closed circular state (where the energy of deformation is at a minimum and the writhing number is zero). For example, for the DNA's of both polyoma virus and the monkey virus SV40, ΔLk is approximately −25. (It is assumed that the linking number is measured with both the relaxed and the supercoiled DNA in solution under standard conditions of temperature, salinity and so on.)

It can now be understood why a deficit in the linking number of a molecule of DNA causes the molecule to supercoil. A linear molecule of DNA in solution normally assumes a form known as the *B* configuration, in which the nucleotide bases are approximately perpendicular to the helical axis with 3.4 angstrom units between them and in which there are about 10 base pairs for each turn of the double helix. This is a configuration of minimum energy, and if the molecule is bent or twisted, its energy is increased. If a long molecule is simply circular, however, the diameter of the circle is large compared with the thickness of the double helix. Hence the curvature of the molecule is small and its energy is increased only slightly. As a result nicked circular molecules such as component II of polyoma-virus DNA hardly depart from the *B* configuration.

The situation is quite different for a closed circular molecule with a deficit in linking number. To satisfy the condition that the value of Lk be less than that of a relaxed molecule (say 475 rather than 500) the double helix would have to be untwisted, a transformation that would substantially increase the deformation energy of the molecule. By supercoiling, however, the closed circular molecule minimizes the amount by which it departs from the *B* configuration.

More precisely, as the analysis of the ribbon model revealed, one way that underwound DNA can reduce its deformation energy is by writhing. Since writhing and twist are interconvertible, it is apparent that by changing the extent of writhing it is possible to minimize the twist of a molecule, thereby minimizing the twisting component of its deformation energy. On the other hand, writhing always introduces some curvature, and so it increases the bending contribution to the energy of the molecule. Therefore the supercoiled configuration that the underwound DNA molecule assumes is one that minimizes twist while introducing the smallest possible amount of bending.

Thus when the axis of DNA is made to writhe, the double helix responds by twisting and/or bending. The forms that this supercoiling can take range from a left-handed solenoidal superhelix to a right-handed interwound superhelix, although in nature the interwound



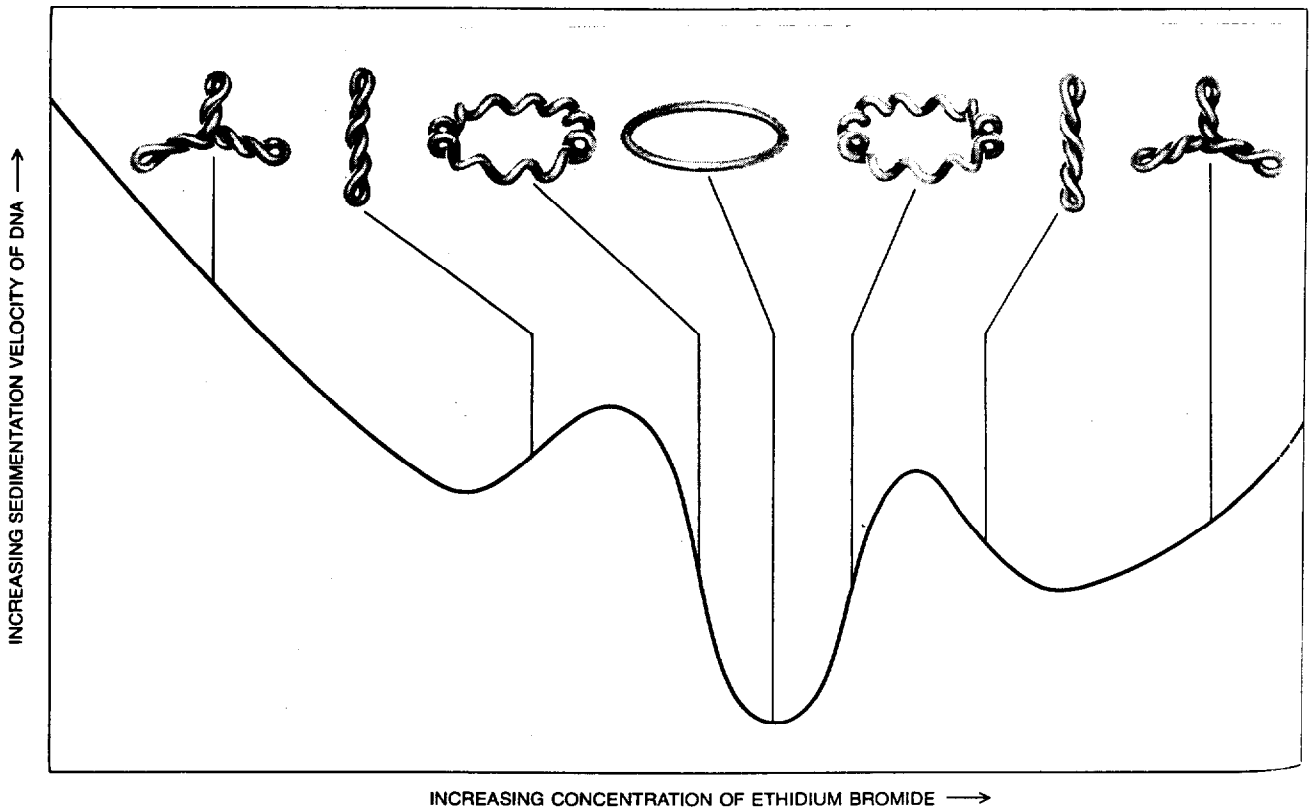
RUBBER TUBE that has been coiled around a cylinder with its ends joined in such a way that all twist is relieved (*left*) jumps into an interwound helix coiled in the opposite direction when the cylinder is removed (*right*). If there are N right-handed turns in the first configuration, and if the pitch angle α_1 at which each helical turn inclines away from the horizontal is small, then in the second configuration there will be approximately $N/2$ turns going up and $N/2$ going down and the new pitch angle α_2 will be large. (This will be only roughly true.) An examination of the changes in the values of the linking number, the total twist and writhing number that accompany this transformation explains why most naturally occurring DNA, having a deficit in linking number, supercoils into interwound helix that like double helix is right-handed.

configuration is usually preferred. The trade-off between writhing and twist that results in the favoring of the interwound form may be made clearer by a mechanical example. Consider a piece of thick, flexible rubber tubing wound into a left-handed coil around the surface of a cylinder. Assume that as the tubing is coiled it is allowed to twist freely about its own axis, so that its twisting energy is zero. Its ends are then brought together (in such a way that any strain due to twisting is relieved) and permanently joined. If at this point all restraint on the tubing is removed, however, it will jump into an interwound superhelix that is right-handed [*see illustration on preceding page*]. Why?

Without undertaking a detailed analysis of the mechanics of this situation it is possible to gain a general understanding of how a left-handed superhelical coil that has little or no twist but is highly writhed is transformed into a right-handed interwound superhelix by considering the linking number, the total twist and the writhing number of the two configurations. (To visualize the meaning of linking number in this context imagine a pair of parallel lines drawn on opposite sides of the unwound relaxed tubing.) Remember that in the first configuration the total twist was deliberately made to be zero, and so (ignoring end effects) the linking number of the tubing must be equal to the writhing number. In the example of the helical ribbon wound around a cylinder we observed that the writhing number for such a configuration is given by the formula $Wr = -N_1(1 - \sin\alpha_1)$, where N_1 is the number of coils in the tubing and α_1 is their pitch angle. (The negative sign arises because the coils are left-handed.) Hence for the first configuration the linking number equals $-N_1(1 - \sin\alpha_1)$. (Because we are neglecting end effects this number will not usually be an integer.)

Turning now to the second configuration of the rubber tubing, the right-handed interwound superhelix, we shall call the total number of right-handed turns (up plus down) of the helix N_2 and its pitch angle α_2. Then (once again neglecting end effects) it can be shown that for such an interwound superhelix the writhing number Wr equals $-N_2\sin\alpha_2$. Note that the left-handed superhelical coil was transformed into the interwound superhelix without breaking the rubber tube or separating its ends. Therefore the linking number of the second configuration must be the same as the linking number of the first, that is, Lk must be equal to $-N_1(1 - \sin\alpha_1)$. Substituting these values of Wr and Lk into the equation $Wr = Lk - Tw$ yields $-N_2\sin\alpha_2 = -N_1(1 - \sin\alpha_1) - Tw$, or $Tw = -N_1(1 - \sin\alpha_1) + N_2\sin\alpha_2$.

The pitch angle of the first configuration of the rubber tubing α_1 was small, but the pitch angle α_2 of the second configuration was fairly large. Therefore matters can be further simplified by substituting the approximations $\sin\alpha_1 = 0$ and $\sin\alpha_2 = 1$ into the equation, yielding $Tw = -N_1 + N_2$. This simple formula explains why the rubber tube when it was released jumped into an interwound configuration of reversed handedness. To these approximations the twist of the tubing Tw is at a minimum when the turns in the second configuration are equal in number and opposite in handedness to those in the first configuration. Moreover, the curvature of the second



SUPERCOILS ARE RELAXED when the dye ethidium bromide, which has a planar molecular structure, is added to supercoiled DNA in solution, as is demonstrated by this graph showing how the sedimentation velocity of DNA changes as the dye is added. The sedimentation velocity is the rate at which the molecules of DNA move through the solvent when the solution is subjected to the strong gravitational field generated by an ultracentrifuge: more highly supercoiled molecules are more compact, and so they have a greater sedimentation velocity. Increasing the concentration of ethidium bromide serves to gradually decrease the writhing of the DNA molecules, not by reducing the deficit in their linking number (that can be done only by creating nicks in their polynucleotide chains) but by reducing their twist. Molecules of the dye are intercalated between the base pairs of the DNA, creating a local untwisting of the double helix. The graph shows that with a sufficiently high concentration of dye the DNA molecules become fully relaxed, and that if more dye is added, the DNA begins to supercoil in the opposite direction. Changes in the structure of DNA shown here are confirmed by electron microscopy.

configuration is clearly less than that of the first. Therefore when the twist is small, the interwound superhelix will have an energy of deformation much lower than that of the superhelical coil, and so the rubber tube will of course assume the interwound form. (To apply this argument to DNA we should have to paint the double helix onto the unwound relaxed rubber tubing, but that would not affect the energy calculation.)

To make a detailed analysis of the supercoiling of real DNA requires a fairly precise knowledge of all its elastic constants, and it is necessary to take into account not only end effects but also such matters as charge repulsion and thermal motion. The preceding arguments show quite clearly, however, why a closed circular DNA molecule, with a deficit in its linking number (or as is sometimes the case an excess) will writhe into a supercoil. In nature most closed circular DNA is negatively supercoiled, that is, the supercoiling results from a deficit in linking number. The analysis of the rubber-tube model explains why such molecules can be expected to assume the configuration of an interwound superhelix whose handedness, rather surprisingly, is the same as that of the double helix, namely right-handed.

The backbones of a double helix have enough flexibility so that if in a supercoiled molecule one of them is nicked, the other can rotate about it to put the molecule in a relaxed configuration. The un-nicked molecule cannot lose its supercoiling in this way without one intact backbone's passing through the other, which is a physical impossibility. Given a naturally supercoiled molecule of DNA with a deficit in its linking number, however, is it possible for its writhing to be reduced without one of its backbones' being nicked? If the backbones must remain intact, then the linking number cannot change, and since the writhing number is negative in this case, the only way for the magnitude of the molecule's writhing to be reduced is for its total twist to be reduced.

Adding ethidium bromide, a dye with a planar molecular structure, to a solution of DNA serves this purpose. Leonard S. Lerman of the State University of New York at Albany was the first to show that molecules of this type are intercalated between the base pairs of the double helix. In supercoiled DNA such insertion causes a local untwisting of the double helix. Indeed, a sufficiently high concentration of ethidium bromide will completely relax the molecule.

More precisely, this critical concentration will reduce the total twist of the molecule. For example, in the presence of the dye a linear piece of polyoma-virus DNA would have a total twist of 475 rather than 500. Therefore a closed circular molecule of that DNA with a linking number of 475 would in the presence of this critical amount of dye have a total twist equal to its linking number, so that its writhing number would be zero. The axis of the molecule would cease writhing. Furthermore, if the concentration of ethidium bromide is increased past the critical point, the total twist will be still further reduced, causing the closed circular molecule to writhe into an oppositely oriented supercoil.

These effects can be observed directly by studying the sedimentation velocity of DNA molecules, as was first shown by Lionel V. Crawford of the Imperial Cancer Research Institute in Britain and Michael J. Waring of the University of Cambridge and independently by Vinograd and one of us (Bauer). Adding ethidium bromide to closed circular polyoma-virus DNA in solution causes the molecules of DNA to writhe less, so that they become less compact and sediment slower. As more ethidium bromide is added their sedimentation velocity reaches a minimum; then as higher concentrations of the dye cause the molecules to supercoil in the opposite direction and become more compact the sedimentation velocity begins to rise again.

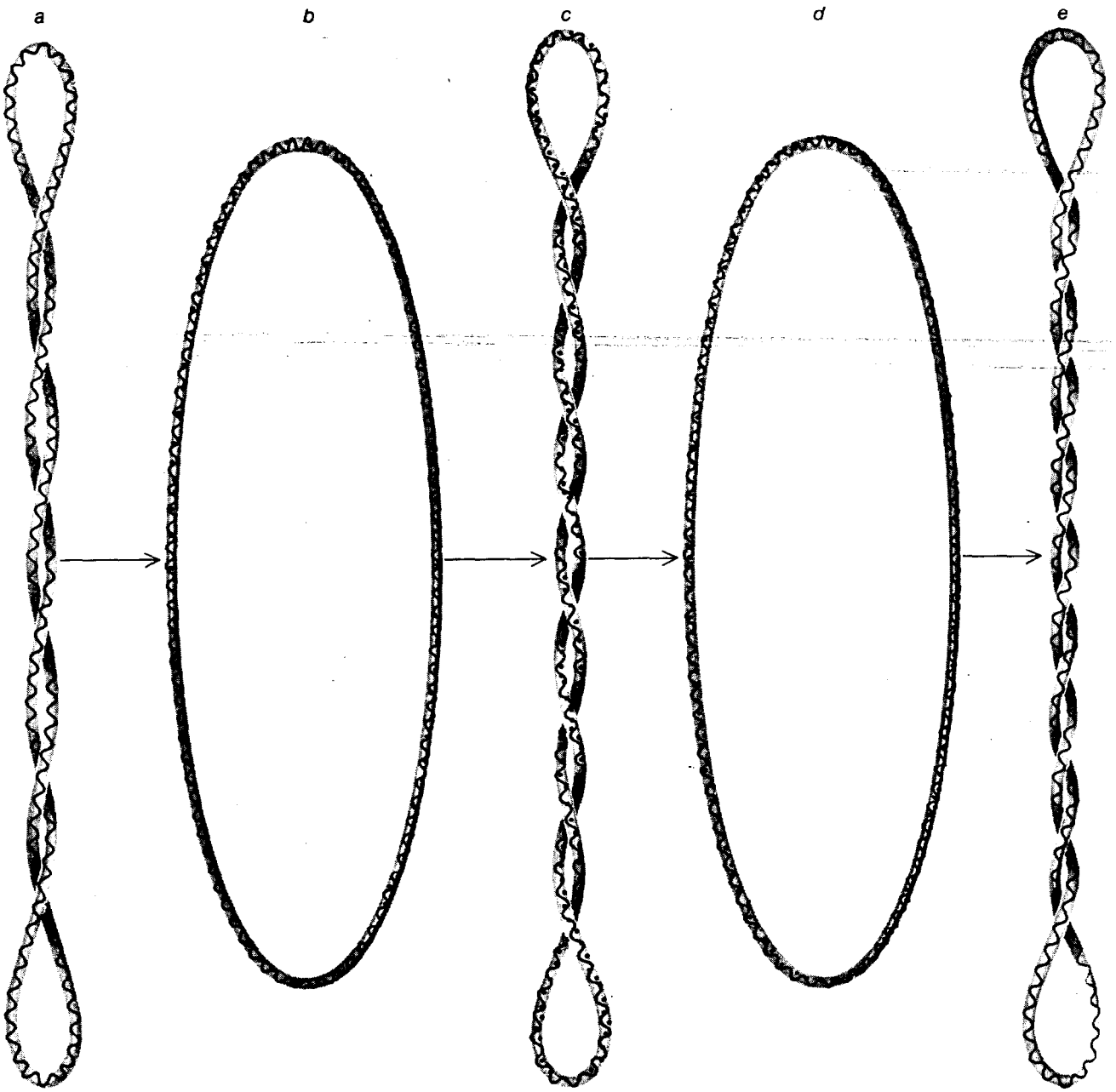
To create DNA molecules with different degrees of writhing it would be convenient to be able to make a nick in one of the backbones of a DNA double helix, relax the molecule by a few turns and then close the nick. Astonishingly, enzymes have been identified that do just that. The first of them, an enzyme known as ω-protein, was discovered in *Escherichia coli* by James C. Wang of Harvard University in 1971. Similar enzymes have now been discovered in a variety of sources, including other bacteria, the vaccinia virus and the nucleus and mitochondria of animal cells. These nicking-closing enzymes, which are also called topoisomerases, generally require no energy source to function. They always act to reduce the supercoiling of a DNA molecule and thereby lower its energy.

Nicking-closing enzymes have proved to be invaluable tools for studying the physical chemistry of DNA. To begin with, adding a topoisomerase to a solution of supercoiled DNA gradually reduces the supercoiling until all the DNA molecules are in or near the relaxed state, but much more ingenious applications have been devised for these enzymes. For example, suppose a high concentration of ethidium bromide is added to closed circular DNA, causing it to supercoil in the direction opposite to its usual one. If a topoisomerase is then added as well, the molecules in the solution will become relaxed, but because of the large amount of intercalated ethidium bromide they will be considerably more underwound than they are in their natural state. Now, if first the

topoisomerase and then the dye is removed, the molecules will compensate for the additional reduction of the linking number by supercoiling even more strongly in the usual direction. In this way super-supercoiled molecules can be created, molecules that have many more superhelical twists than are usually present. This type of energy storage and transfer may be central to the role of supercoiling in the cell. We can even conceive of a chemical engine that by going through a similar cycle of operations turns chemical energy into work.

It is also interesting to note that Martin Gellart and his colleagues at the National Institutes of Health have found an enzyme in the bacterium *E. coli* that given an energy source such as adenosine triphosphate (ATP) will reduce the linking number of relaxed circular DNA, thereby increasing its writhing number and making it supercoil. The activity of an enzyme of this type, called a gyrase, is essentially opposite to that of the nicking-closing enzymes. So far gyrases have been found in a variety of microorganisms but not in any higher organisms.

Nicholas R. Cozzarelli and his colleagues at the University of Chicago have recently shown that a gyrase seems to act by bringing two segments of the DNA molecule close together. The en-



SUPER-SUPERCOILED MOLECULES OF DNA, with more superhelical turns than are normally present, can be created by employing ethidium bromide (*colored dots*) in conjunction with a nicking-closing enzyme. An enzyme of this type acts to reduce the supercoiling of a DNA molecule by making a temporary nick in one of its sugar-phosphate backbones, relaxing the molecule and then reclosing the nick. Therefore if enough ethidium bromide is added to supercoiled DNA in solution (*a*) for the DNA to become first completely relaxed (*b*) and then supercoiled in the opposite direction (*c*), the addition of the nicking-closing enzyme will bring the DNA back to the relaxed state (*d*). A substantial amount of dye will be bound to the DNA, however, so that the DNA will be more underwound than is usual. Hence if the enzyme and then the dye are removed, the DNA will supercoil more strongly than usual in its original direction (e)

zyme then cuts both of the backbones of one of the segments and passes the other intact segment through the resulting gap before it rejoins the originally cut backbones. It is easy to show that such a process would alter the linking number of the DNA in increments of two rather than one, and that is exactly what has been observed experimentally.

How is supercoiling measured? One experimental method is based on a discovery made by Walter Keller of the University of Heidelberg in 1974. Keller observed that if supercoiled molecules with even slightly different linking numbers are moved through an agarose gel by electrophoresis, they do not all travel at the same rate. The more compact molecules travel faster, and so a pattern of discrete bands is created in the gel. Moreover, as was later shown by Vinograd, if one of these bands of molecules is cut out of the gel and rerun, the result is again a single band in exactly the same location. In fact, as was shown later by Wang and one of us (Bauer), even if the band is heated to near melting (so that the bases in the double helixes become unpaired) and then is slowly cooled (so that the double helixes are rejoined), it will still form a single band in the same location.

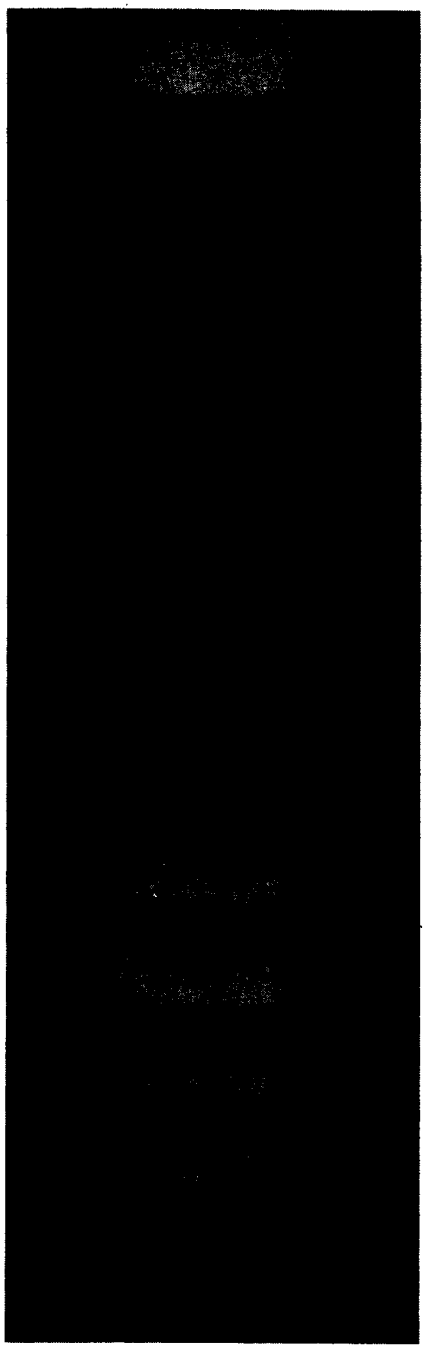
Hence the molecules in adjacent bands differ by a quantity that survives heating and cooling and, since the bands are discrete, is quantized. This quantity can only be the linking number. The only reasonable interpretation is that the linking numbers of the molecules in adjacent bands on the gel differ by a value of 1. Therefore when a population of DNA that includes molecules in all the states from relaxed to supercoiled is run on the gel, ΔLk, the difference in linking number that determines supercoiling, can be evaluated directly by counting the bands. This experimentally determined value for ΔLk can be put to good use. For example, it can be applied to calibrate the untwisting effect of ethidium bromide on the double helix. It turns out that every intercalated molecule of ethidium bromide produces a local untwisting of about 26 degrees, and so ethidium bromide can be utilized as a subsidiary standard for measuring the value of ΔLk for any supercoiled molecule. (It should also be noted that the striking effect of the enzyme gyrase on the linking number of DNA can be observed clearly in gel experiments.)

One of the most important results to emerge so far from the study of supercoiled DNA is the direct proof that the two backbones of DNA do indeed wind around each other at intervals of about 10 base pairs. Here the most elegant experiments are those that have been done by Wang. He compared similar circular DNA molecules constructed by the methods of genetic engineering, each having about 5,000 base pairs. The exact length of these molecules was not known, but the difference in length between any two of them was known precisely. Some of them differed by only a single base pair, others by as many as 400 pairs.

The differences in the extent of supercoiling observed when these different-sized molecules were relaxed by means of a nicking-closing enzyme and then run side by side on a gel enabled Wang to deduce that the number of base pairs for each turn of the double helix must be 10.4 plus or minus .1. His experiment was a pretty demonstration of the precision and sophistication of modern techniques of molecular biology. Moreover, several models for DNA have recently been proposed in which the two polynucleotide chains do not coil around each other to form a double helix but instead lie side by side over most of their length, having only a few helical turns. Wang's results, together with those from related experiments, show that these new models must be incorrect. This topological argument is very powerful in that it eliminates all models of the side-by-side type, regardless of their molecular detail.

How does supercoiling arise? In the SV40 virus and the polyoma virus the DNA is supercoiled because it is usually not naked in the nucleus of the cell. During replication the double helix is wound on nucleosomes (beads of protein consisting of eight histone molecules with one or two associated molecules). When the DNA molecule slips off this supporting structure, it supercoils (in much the same way that the coiled rubber tubing discussed above does). The fact that the linking number is always reduced in naturally occurring DNA implies that when DNA winds around the nucleosomes in a solenoidal coil, the coil is left-handed. Uncertainties in the available data and arguments about their interpretation make it difficult to determine the exact number of supercoils that are generated by each nucleosome, but it appears likely that the number will turn out to be between 1 and 2. Nucleosomes are found in association with DNA in all higher organisms; in lower organisms the origins of supercoiling are not completely clear.

Much work remains to be done before all the implications of supercoiling in DNA can be understood. In particular the enzymes that are related to supercoiling (topoisomerases and gyrases) and the role of supercoiling in the replication of DNA are currently the subject of intense research efforts. It is clear, however, that the present knowledge of supercoiling, both experimental and theoretical, provides a sound basis for the investigation of this surprising manifestation of the double helix.



EXPERIMENTAL TECHNIQUE for measuring supercoiling involves moving a population of closed circular DNA molecules through an agarose gel by electrophoresis, in this case from top to bottom. Because the more highly supercoiled molecules are more compact they move faster through the gel. As a result a set of discrete bands of molecules is created such that the molecules in each band all have the same linking number and the numbers associated with adjacent bands differ by a value of 1. (The DNA shown here has been stained with a fluorescent dye, and so the bands containing more molecules are brighter.) If the population includes molecules with various degrees of supercoiling, as is shown here, counting the resulting bands provides a direct measure of the deficit in linking number that determines degree of supercoiling. Uppermost band in gel contains nicked DNA.